PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/82, 15/29, A01H 5/00

A2

(11) International Publication Number: **WO 97/20058**

(43) International Publication Date: 5 June 1997 (05.06.97)

(21) International Application Number: PCT/HU96/00070

(22) International Filing Date: 22 November 1996 (22.11.96)

(30) Priority Data:
P 95 03352 24 November 1995 (24.11.95) HU

(71) Applicants *(for all designated States except US)*: BAY ZOLTÁN ALKALMAZOTT KUTATÁSI ALAPÍTVÁNY [HU/HU]; Fehérvári út 130, H-1116 Budapest (HU). MTA SZEGEDI BIOLÓGIAI KÖZPONT NÖVÉNYBIOLÓGIAI INTÉZET [HU/HU]; Temesvári krt. 62, H-6701 Szeged (HU).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: KAPROS, Tamás [HU/US]; 4408 Pennsylvania Avenue 3, Kansas City, MO 64110-2499 (US). DUDITS, Dénes [HU/HU]; Bérkert u. 36-3, H-6726 Szeged (HU). GYÖRGYEY, János [HU/HU]; Alsó Kikötő sor. 5 VIII.48, H-6726 Szeged (HU). MAI, Antal [HU/HU]; Szamos u. 1-A, V.17, H-6723 Szeged (HU). KELEMEN, Zsolt [HU/HU]; Retek. u. 5 VI/23, H-6723 Szeged (HU).

(74) Agent: DANUBIA; P.O. Box 198, H-1368 Budapest (HU).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: PLANT GENE EXPRESSION VECTOR FAMILY BASED ON THE REGULATORY DNA SEQUENCES OF AN ALFALFA H3 HISTONE GENE VARIANT (MsH3g1)

(57) Abstract

Disclosed herein are nucleic acid molecules comprising sequence regions participating in the regulation of the expression of the H3g1 gene of alfalfa (*Medicago sativa*). A plant gene expression vector family based on the regulatory DNA sequences of said histone gene variant (MsH3g1) is also disclosed. The invention also provides transformed cells, transgenic plants and parts thereof comprising the nucleic acid sequences of the invention.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

Plant gene expression vector family based on the regulatory DNA sequences of an alfalfa H3 histone gene variant (MsH3g1)

The present invention relates to nucleic acid molecules comprising sequence regions participating in the regulation of the expression of the H3g1 gene of alfalfa (*Medicago sativa*) and a gene expression vector family based on the nucleic acid molecules of the invention. More specifically, the invention relates to nucleic acid molecules comprising sequences presented on Fig. 1. The invention also provides transformed cells, transgenic plants and parts thereof comprising the nucleic acid sequences of the invention.

The present invention is useful in expressing foreign genes in different plants and in altering the measure or spatial and temporal pattern of the expression of different endogenous plant genes.

With respect to the present specification and claims, the foregoing technical terms will be used in accordance with the below given definitions. With regard to the interpretation of the present invention, it shall be understood that the below defined terms are used in accordance with the given definitions even if said definitions might not be in perfect harmony with the usual interpretation of said technical term.

A *"homologue"* or a *"variant"* of a nucleic acid sequence is defined as a sequence that is at least 50 % identical to the sequence in question.

A *"functional variant"* of a sequence is every sequence having the same type of biological activity even if the measure of the biological activity of the functional variant is significantly different from that of the original sequence (e.g., the transcriptional activity of the functional variant of a promoter can either be smaller or larger than that of the original promoter).

A nucleic acid molecule is regarded *"hybridizable"* with another nucleic acid molecule if it can specifically be bound to the other molecule (i.e., the binding can give rise to a signal that is distinguishable from the background noise and from the signal caused by the aspecific binding of any random sequenced nucleic acid molecule).

A regulatory sequence is *"operatively linked"* to structural gene within a DNA construct if the regulatory sequence is able to influence the expression rate or manner of said structural gene under conditions suitable for the expression of said structural gene and for the functioning of said regulatory sequence.

The first scientific papers disclosing so called transgenic plants - prepared by transformation and comprising foreign genes - were published in the early 1980's [Fraley, R.T. et al., Proc. Natl. Acad. Sci. USA 80: 4803 (1983)]. Since that time transformation methods has been elaborated for almost all economically important industrial crops [for review see Hinchee et al., Plant Cell and Tissue Culture, pages 231-270, ed. Vasil, J.K. and Thorpe, T.A., Kluwer Academic (1994)]. Transgenic techniques on one hand became

generally applied research tools in the field of experimental plant biology and, on the other hand, transformed plants are now getting used worldwide in plant breeding systems and in the seed-corn industry. In 1993, statistics account for 320 announced field experiments establishing the widespread application of transgenic plants. There is a commercially available transgenic tomato line in the USA, that is advantageously marketable because of its better taste and applicability for long term storage.

There are a lot of factors influencing the advantageous applications of transgenic plants, one of the most important of which - with respect to the establishment of new technologies - is the appropriate choice of the so called "agronomic" genes (genes causing positive effects) and the regulatory elements directing their expression (e.g., promoters, introns and 3′ regulatory regions). A very narrow range of promoters has been used so far for the expression of foreign genes in plants. The cauliflower mosaic virus derived CaMv35S promoter is most often used for the construction of plant expression vectors. With regard to the limitations of such expression systems there is a rapidly growing need for the preparation of further vector systems based on different promoters. In a wide range of possible applications it would be advantageous to use so called constitutive vector constructions ensuring high expression rate in every cells and tissue types of the transformed plant. It would also be preferred to use plant promoters

instead of virus derived promoters to avoid possible biohazard.

The concentration of a given gene product in a cell is influenced by a multilevel system of molecular mechanisms. Highly important factors of such mechanisms are the speed of transcription, the term of half life of the produced mRNA molecules and the stability of said molecules. These parameters are mainly dependent on the characteristics of the 5′ promoter region and can also be influenced by the introns present in the coding region and the 3′ non translating regions. Accordingly, the above regulatory elements of a plant gene of interest can be used for constructing transformation vectors ensuring high expression levels of agronomic genes.

For the isolation of constitutive promoters, it is advisable to consider genes participating in basic functions of the cells. One such type of genes can be the histone gene family. The present inventors has conducted intensive research in the field plant histone genes since the mid 1980′s, primarily using experimental systems based on alfalfa. Our publication, in which we have demonstrated the presence of H3 histone gene variants in alfalfa, was published in 1988 (Wu, S.C., Bögre, L., Vincze, E., Kiss, G.B. and Dudits, D., Plant Mol. Biol. 11: 641). In this paper - by applying cDNA clones - we have demonstrated the presence of a cell cycle dependent (H3.1) and a constitutive gene variant (H3.2). We have isolated the genomic clone of the H3.1 variant from a gene bank and characterized the cell

cycle dependent functioning of its promoter [Kapros, T. et al., Plant Physiol. 98: 621 (1992)] . We have also isolated the genomic counterpart of the constitutive cDNA variant and characterized the regulatory elements of this gene that make us possible to use said regulatory elements for the development of new transformation vector molecules.

It is, thus, an object of the present invention to construct a wide host range constitutive plant gene expression vector family based on the regulatory regions of the isolated MsH3g1 gene variant.

The object of the invention was achieved by first determining the whole nucleotide sequence of the MsH3g1 gene. Based on the gained sequence information it became possible to determine the location of the 5′ promoter region (from -482 to -1 bp with respect to the transcription initiation site; nucleotides 1-482 of the sequence shown on Fig. 1A), of three introns (nucleotides 555-668, 746-962, 1053-1174; Fig. 1B) and of the 3′ non translating region (nucleotides 1346-1676; Fig. 1C).

The above regulatory regions are suitable - alone or in combination - for constructing novel vector molecules making possible the functioning of foreign genes in plant cells or the alteration of the spatial and/or temporal expression pattern of endogenous plant genes.

The present invention provides single or double stranded DNA molecules comprising - wholly or in part - the sequence regions of the H3g1 gene of alfalfa (*Medicago sativa*) shown on Fig. 1A (nucleotides 1-482), Fig. 1B

(nucleotides 555-668, 746-962 and 1053-1174) and/or Fig. 1C (nucleotides 1346-1676), or their homologues or functional variants.

The nucleic acid molecules of the invention are advantageously at least 80 % homologue to the sequences shown on Fig. 1 or parts thereof.

The invention also concerns nucleic acid molecules hybridizable to the nucleic acid molecules of the invention.

The invention further provides transformation vectors comprising the nucleic acid molecules according to the invention, cells and transgenic plants transformed with such vectors, and reproducible parts of said transgenic plants.

The invention is further illustrated by the attached figures the short description of which is as follows:

Fig. 1 shows the regulatory regions of the MsH3g1 gene (Fig. 1A depicts the sequence of the promoter region; Fig. 1B shows the sequences of introns I, II and III; and Fig 1C shows the 3' non translating region).

Fig. 2 shows the schemes of vector constructions prepared using the regulatory regions according to the invention (see also Ex. 2).

Fig. 3 shows the schemes of further vector constructions according to the invention (see also Ex. 2).

The subject of the invention will be further illustrated by the experimental examples described below, however, the scope of the invention will by no means be limited to the specific embodiments described in the examples.

Example 1

Determination of the nucleic acid sequence of the MsH3g1 gene isolated from alfalfa (*Medicago sativa*), identification of functional elements of the gene

We have isolated a clone comprising the MsH3g1 gene from phage λ clones (Alfalfa Genomic Library, supplied by Clontech Laboratories, Inc.) carrying alfalfa genomic DNA applying a colony hybridization method known in the art [Sambrook et al., Molecular Cloning, A Laboratory Manual, 2. ed., Cold Spring Harbor, N.Y. (1989)]. A PvuII/HindIII fragment of the previously isolated pH3c11 cDNA clone [Wu, S.C. et al., Nucleic Acids. Res. 17: 3057 (1989)] was employed as the hybridization probe. Sequencing of the isolated H3 histone gene was also done using techniques known in the art [Schuwmann, R., et al., Biotechniques 10: 185 (1991); Seto, D., Nucleic Acids. Res. 18: 5905 (1990); Manfioletti, G. et al., Nucleic Acids. Res. 16. 2873 (1988)]. The nucleotide sequence of the 5′ promoter region of the isolated alfalfa H3 histone gene - that is located between nucleotides -482 and -1 - is shown in Fig. 1 (the numbering of this sequence on the figure is 1-482). This DNA segment comprises all the functional elements needed for the expression of a gene of interest. The isolated alfalfa histone gene comprises three non coding intron regions shown on Fig. 1B. The 3′ non coding region is comprised of 230 nucleotides (Fig. 1C).

Example 2

Preparation of a transformation vector family based on the functional regulatory elements of the MsH3g1 gene

The vector molecules shown on Fig. 2 and Fig. 3 were prepared using DNA manipulation methods generally known in the art [Sambrook et al.: Molecular Cloning, A Laboratory Manual, 2. ed., Cold Spring Harbor, N.Y. (1989)] . The following fragments of the MsH3g1 genomic clone were used to construct said vector molecules: the promoter region flanked by 5′ AccI and 3′ NcoI restriction sites, the intron I region flanked by NsiI/NcoI sites and the transcription termination region flanked by SacI/NotI restriction cleavage sites. The applied multiple cloning sites (MCS) region was derived from pBluescript II SK+ (Stratagene; GenBank #X52328[SK(+)] , designated as plasmid pBSK on the figures). For the construction of vectors pHEX-N and pHEX110, we have also used the termination signal of the bacterial nopaline synthase gene [Depicker et al., J. Mol. Appl. Genet. 1: 561 (1982)] .

We summarize the construction of the above vector molecules in the following steps.

1. The promoter region was cut from plasmid construction pHEX-N (Fig. 2) by AccI and NcoI restriction cleavage, overhangs were filled in with Klenow enzyme (*E. coli* DNA polymerase I, Large Fragment, Biolab′s), then the isolated promoter fragment was ligated into plasmid pBluescript II SK+ previously cut with enzymes XbaI and BamHI and simi-

larly filled in with Klenow enzyme, resulting in plasmid pBH.

2. The NOS terminator was isolated on a KpnI/SacI fragment and inserted into the KpnI/SacI sites of plasmid pBH yielding plasmid pBHN. Plasmid pBHN comprises the following construction: - NotI - XbaI - H3.2 promoter - NcoI - pBluescript II MCS (from BamHI to KpnI) - SacI - NOS terminator - NotI -.

3. Plasmid pBluescript II SK+ was cleaved with enzymes KpnI and SacI and a synthetic NotI linker was inserted into the cleaved plasmid, resulting in plasmid pBNot. The sequence of the inserted NotI linker was the following:

upper strand: 5′ CCCGCGGCCGCCCCAGCT 3′

lower strand: 3′ CTAGGGGCGCCGGCGGGG 5′ .

4. Plasmid pBHN was cut with NotI and the isolated fragment comprising the - H3.2 - promoter - MCS - NOS terminator - construction was ligated into the NotI site of plasmid pBNot, resulting in plasmid pHEX-N2.

5. Plasmid pHEX-N2 was digested with enzymes NcoI and SacI and a NcoI/SacI fragment of plasmid pLP140 [Szczyglowski, K. et al., The Plant Cell 6: 317 (1994)] comprising the coding sequence of the GUS gene was inserted into the cleaved plasmid, yielding plasmid pHEX-N::GUS.

6. Intron I of the H3.2 genomic clone was amplified by PCR and the PCR product was introduced to the NcoI site of plasmid pBHN2, resulting in plasmid pHEX110.

7. The GUS coding sequence comprised on a NcoI/SacI fragment was then inserted into the NcoI/SacI sites of plasmid pHEX110, yielding plasmid pHEX110::GUS.

8. The NOS terminator in pHEX110 vector was replaced by the H3.2 terminator amplified by PCR using the pH3.2 genomic clone as template and applying techniques well known in the art, resulting in plasmid pHEX111. The sequences of the primers used in the PCR reaction were as follows:

upstream primer:

5′ GAG CTC TAG GTA GGT AGC ATT CGC GGT GAA CGT GCT 3′

downstream primer:

5′ GCG GCC GCT GTC ACC GAT AGA CAA ACT ACC 3′ .

The PCR product was polished with Klenow fragment of DNA Polymerase I and then it was subcloned into the T4 polymerase-blunted KpnI/SacI restriction sites of plasmid pBluescriptSK+. The SacI/SacI fragment of the resulting plasmid was isolated and inserted into the SacI/SacI sites of plasmid pHEX110 replacing the NOS terminator by the alfalfa histone terminator. This resulting construction was designated pHEX111.

9. The GUS coding region was then inserted to the NcoI/SacI sites of pHEX111, yielding plasmid pHEX111::GUS.

Example 3

Functioning of the pHEX expression vectors in transient experiments performed after transforming alfalfa and maize protoplasts

With the purpose of checking the proper functioning of the transformation vectors shown on Fig. 2, we have also prepared their variants comprising reporter genes. The scheme of vectors carrying the β-glucuronidase (GUS) reporter gene is shown on Fig. 3.

Protoplasts isolated from A2 alfalfa cell suspension [described in Magyar et al., The Plant Journal 4: 151 (1993)] and from HE89 maize cell suspension [Mórocz et al., Theor. Appl. Genet. 80: 721 (1990)] were used in the experiments. To $2x10^6$ protoplasts, 20 μg of plasmid DNA was added and DNA uptake was induced by treatment with 40 % PEG [Oznirulleh, S. et al., Gene transfer to plants, ed. Potrykus and Spangenberg, G., Springer, pages 99-105 (1995)]. Treated protoplasts were cultured in K3-medium [Nagy, J.I. et al., 2. Pflanzenphssiol. 78: 453 (1976)]. For measuring the activity of the β-glucuronidase enzyme, $4x10^5$ protoplasts each was used on the 1th, 2nd and 3rd days of cultivation [Martin, T. et al., GUS Protocols, Academic Press, pages 23-59 (1992)]. In the case of maize cell suspensions, GUS activity values determined on the first day did not significantly differ from background values. Table 1 summarizes specific activity values determined in two parallel experiments using alfalfa protoplast derived cell suspen-

sions originating from two independent transformations (data given are mean values from three parallel measurements). Compared to the activity of the virus derived promoter (CaMV35S::GUS), the functioning and higher promoter activity of the pHEX expression vectors is definitely demonstrated. Especially the vector pHEX110::GUS provided extremely high expression rates. The viral vector used as reference (pIDS211) is described in Stefanov et al., Acta Biologica Hungarica 42: 323 (1991).

Very similar results were obtained using the maize protoplast derived cell suspension. Experimental data are summarized in Table 2 (data given are mean values from three parallel measurements).

The proper functioning of the above vectors has also been demonstrated in tobacco protoplast derived cell suspension (data not shown).

Table 1

Transient GUS activity values determined in alfalfa protoplast derived cell suspension

(pmole MU / mg prot. / min)

Experiment 1

	day 1	**day 2**	**day 3**
control	0	0	0
pHEX-N::GUS	301.4	2457.8	5270.0
pHEX110::GUS	4573.6	18831.7	44434.0
pHEX111::GUS	434.5	5384.6	16028.8
pIDS21 (35S::GUS)	57.8	783.2	565.7

(Table 1, continued)

Experiment 2

	day 1	day 2	day 3
control	1.5	0	0
pHEX-N::GUS	353.8	6287.7	15984.8
pHEX110::GUS	1335.9	23441.0	42570.2
pHEX111::GUS	208.8	5268.4	13891.4
pIDS21 (35S::GUS)	22.3	665.2	1044.0

Table 2

Transient GUS activity values determined in maize protoplast derived cell suspension

(pmole MU / mg prot. / min)

	day 2	day 3
control	0.47	0.49
pHEX-N::GUS	0.46	0.99
pHEX110::GUS	62.51	50.90
pHEX111::GUS	6.64	15.81

Example 4

Stable transformation of pHEX expression vectors into plants

The vectors according to the invention were incorporated into the widely used *Agrobacterium* mediated transformation system and stable tobacco transformants were produced. The description of the employed *Agrobacterium* plasmid designated EHA101 can be found in Hood et al., J. Bacteriol. 168: 1291 (1986). Regeneration of the produced transgenic plants carrying the introduced vectors was also according to the above publication. Briefly, the *Agrobacterium* culture co-cultivated with tobacco leaf platelets was further cultured on plant regeneration medium. After eliminating the bacteria, plants were cultivated in *in vitro* culture and in green house and seeds produced were collected. Proper functioning of the vectors of the invention was established by measuring GUS activity of both primary transformants and progeny seedlings. Production of blue indigo dye was demonstrated in leaf tissues by hystochemical staining according to the method of Jefferson et al. [EMBO J. 6: 3901 (1987)]. A GUS activity value between 1968 and 3891 pmole MU / mg protein / min was measured in leaves of tobacco plants transformed with vector pHEX-N and this result clearly demonstrates the proper functioning of the vectors according to the invention in transgenic plants. In transgenic plants carrying vector construction pHEX110::GUS the measured GUS activity value was between 1598 and 4634 pmole MU / mg protein / min.

CLAIMS:

1. Single or double stranded nucleic acid molecule, comprising - wholly or in part - the sequence regions of the alfalfa (*Medicago sativa*) H3g1 gene presented on Fig. 1A, Fig. 1B and/or Fig. 1C, homologues or functional variants thereof.

2. Nucleic acid molecule according to claim 1, comprising a sequence that is at least 80 % homologous to the sequences or any parts thereof presented in Fig. 1.

3. Nucleic acid molecule according to claim 1, comprising a sequence - or any part thereof - that is identical with or at least 80 % homologous to the nucleotide sequence presented in Fig. 1A (nucleotides 1-482).

4. Nucleic acid molecule according to claim 1, comprising a sequence - or any part thereof - that is identical with or at least 80 % homologous to the nucleotide sequences presented in Fig. 1B (nucleotides 555-668, 746-962 and 1053-1174).

5. Nucleic acid molecule according to claim 1, comprising a sequence - or any part thereof - that is identical with or at least 80 % homologous to the nucleotide sequence presented in Fig. 1C (nucleotides 1346-1676).

6. Nucleic acid molecule hybridizable with a nucleic acid molecule according to any one of claims 1-5.

7. Transformation vector according to any one of claims 1-5, comprising sequences presented in Fig. 1A, 1B and/or 1C - parts, homologues or functional variants thereof - operatively linked to a structural gene.

8. Transformation vector according to claim 7, suitable for transforming plant cells.

9. Transgenic plant or any reproducible part thereof comprising a nucleic acid molecule according to any one of claims 1-7.

10. Plant propagation material prepared from or using a plant according to claim 9.

A

```
   1  ACTTNACTAA CGGAGTCTGC ATTTAGGTAC TAAAATGACT AATATAGTCT
  51  ACATTCAGGG ACTATTTTGC AATTTACCTG CATTCAGGGA CTAAAGTGAC
 101  GACTTCTTTC CTATTCAGAG ACTAAAGTGA CCAATCTCTC AAAATGAAGA
 151  TATTTTGTTT TGTTTTGGCG AGATAAGTTG CACACGATTT ACACTCACAA
 201  AAGAAACACA AATTGTCCAC GCTGGCAATC CGCAACTTTA CAAACCAACC
 251  AATCAGAAAC AAACACACGG ATCGCACTTA ATATTTTCAC TTAAAAAACT
 301  CATCATTACC GTTGAAGCAT TCAAAGTCCA CGATCTCTCT ACTCTAATTA
 351  ACCTTCCTTA ATCATCATTA ACCCTTGCAT ATATAAACAC ACTTCTCTTC
 401  AACAACCCTC ATTACACATT TCTTCTCTTT CGCTAAATCT AATCAATCTT
 451  TCCCTCTCTT CGAGCTTTCT CTCTCCGATT CC  482
```

B

Intron I

```
 555  GTAACC ACCGTTCACC GCCGTAACGG TTTTTTCTTC TTTCTGTTTT
 601  CTTGATCTTA GGGTTTCGTT TTCTTCAATT CGAATTTTTT GATTGATTTC
 651  ATCGATTTTT TGGTTCAG  668
```

Intron II

```
                                                    746  GTAAT
 751  TTCCTTTTCC CCAATTTTTA GGTTTTCGGA GTTTTGCAGT TTCTATTATT
 801  AATTTTTTTT AGGTTTTCGT TGTGTTTTGA ATATTCTATT GAATTTTATG
 851  TTNNGAATTT GAATTTTGGA TTCATAATTT TTAGGAATTT GGAGTTTTGT
 901  ATTCTCGGTT TATGATTTTT AGGTTTTCGT AGTTGTTAAT TTTCAATTGT
 951  TGTGATTTAC AG  962
```

Intron III

```
1053  GGTAAATA TTGTTACTTA GTTTCATAAT TGATTTTGTG TGAAATCTTG
1101  TTCCTCTTTG GTTTATTTAA TTTGTAATTG TTGTTTTTGA TATTTTTCTG
1151  ATTAATTTGC TGTTGTTGTT ATAG 1174
```

C

```
                                                   1346  GGTGG
1351  TGAAGGCGCT TTTAGCGTTA TGGTGGATTA GTATTTTGGA AGGATTTAGG
1401  GTTTTATGAA TTGAATTTTC TTTTATGCGT TGTATAGTTC TGAACCTATA
1451  ATGTTCAATC TTTAACAACA GACATATTTT GGATTATGAT TAGTTTTTTG
1501  CGGACAAATT TGTGATGTAA TTGGTCAATT ACAATTGAAG TCTCTGCAAC
1551  TATTTTACTT ATATCTCCAT TGCTTCCTGA TTTCGTTATG CGCTTTTGAT
1601  GATGCGACTG TGGTTTCTAG CTCTGAATTC AATTTGTGAT GCGTTTCTCT
1651  CTATGGTAGT TTGTCTATCG CTGACA  1676
```

Fig.1

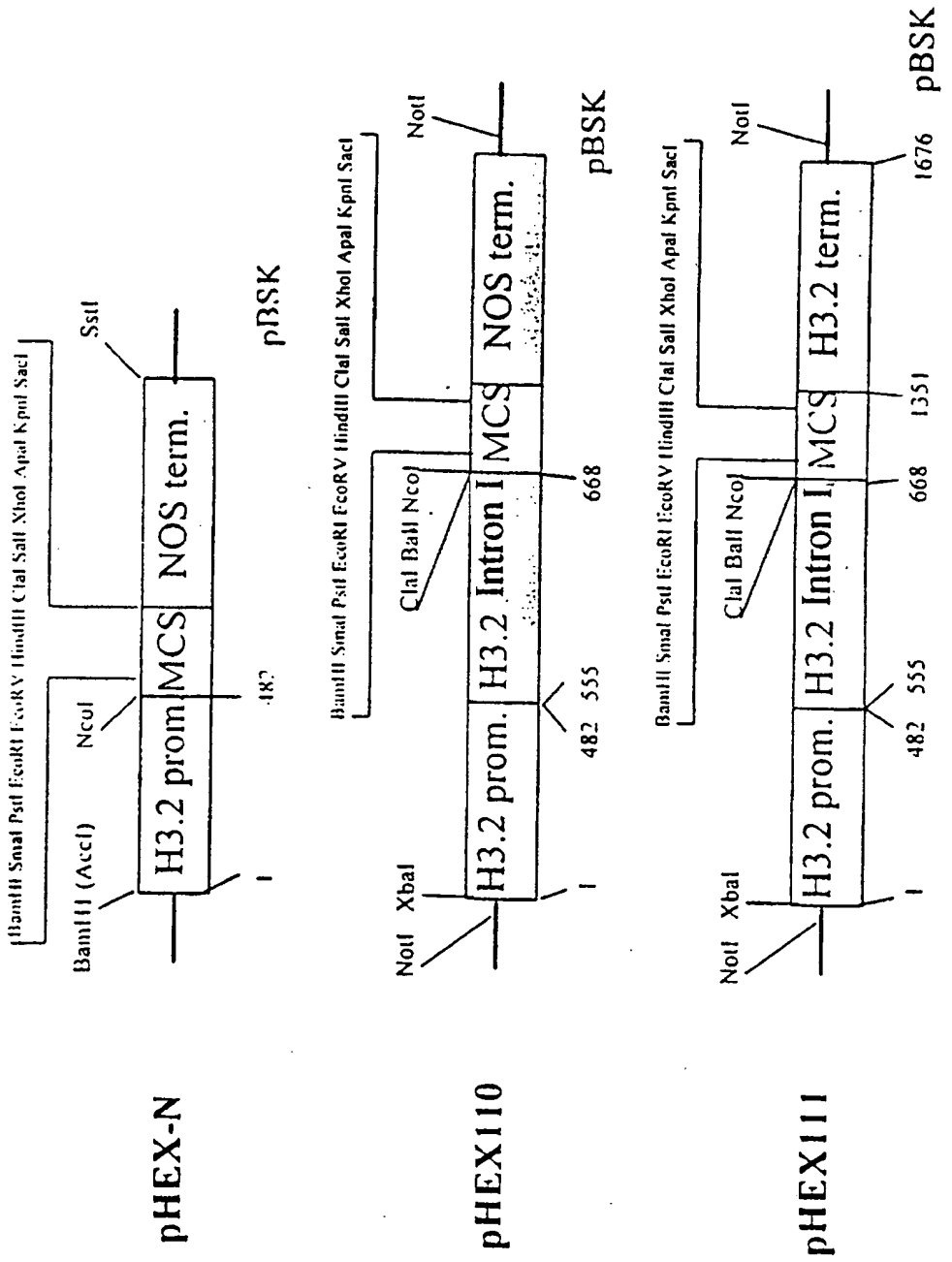
pHEX-N
BamHI Smal PstI EcoRI EcoRV HindIII ClaI SalI XhoI ApaI KpnI SacI
BamHI (AccI)
NcoI
H3.2 prom
MCS
NOS term.
SstI
pBSK
pHEX110
BamHI SmaI PstI EcoRI EcoRV HindIII ClaI SalI XhoI ApaI KpnI SacI
NotI
XbaI
ClaI BalI NcoI
H3.2 prom.
H3.2 Intron I
MCS
NOS term.
NotI
1
482
555
668
pBSK
pHEX111
BamHI SmaI PstI EcoRI EcoRV HindIII ClaI SalI XhoI ApaI KpnI SacI
NotI
XbaI
ClaI BalI NcoI
H3.2 prom.
H3.2 Intron I
MCS
H3.2 term.
NotI
1
482
555
668
1351
1676
pBSK

Fig. 2

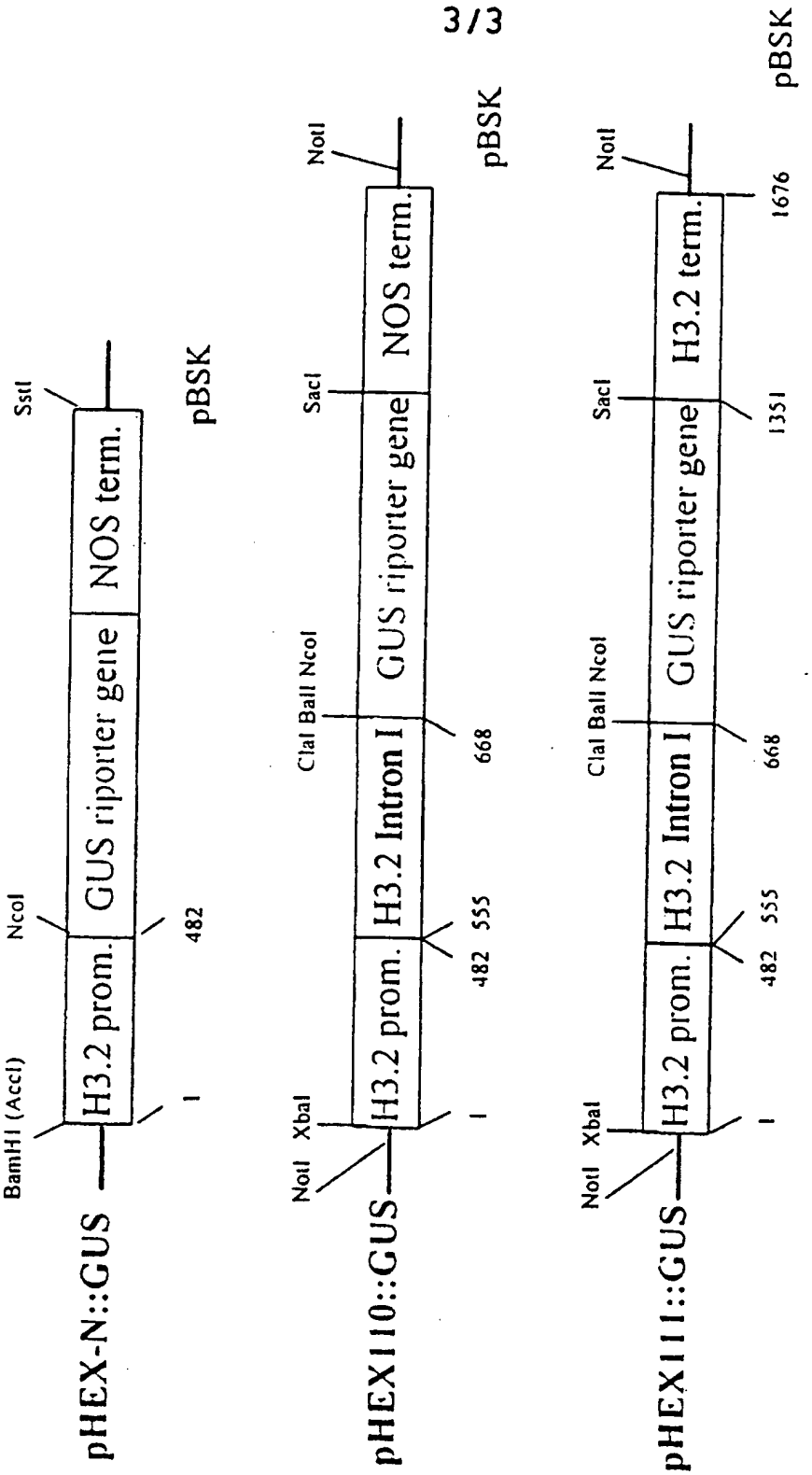
pHEX-N::GUS
BamHI (AccI)
NcoI
SstI
H3.2 prom.
GUS riporter gene
NOS term.
1
482
pBSK
pHEX110::GUS
NotI
XbaI
ClaI
BalI
NcoI
SacI
NotI
H3.2 prom.
H3.2 Intron I
GUS riporter gene
NOS term.
1
482
555
668
pBSK
pHEX111::GUS
NotI
XbaI
ClaI
BalI
NcoI
SacI
NotI
H3.2 prom.
H3.2 Intron I
GUS riporter gene
H3.2 term.
1
482
555
668
1351
1676
pBSK

Fig. 3